UPOV

INTERNATIONAL UNION FOR THE PROTECTION OF NEW VARIETIES OF PLANTS
GENEVA

TECHNICAL WORKING PARTY ON AUTOMATION AND COMPUTER PROGRAMS

Twenty-Fifth Session
Sibiu, Romania, September 3 to 6, 2007

LEAST SIGNIFICANT DIFFERENCE (LSD) AND
OTHER TESTS TO COMPARE MEANS

Document prepared by an expert from the United Kingdom

LEAST SIGNIFICANT DIFFERENCE (LSD) AND
OTHER TESTS TO COMPARE MEANS

Mrs. Sally WATSON
Biometrics Branch
Agri-Food & Biosciences Institute
18a, Newforge Lane
Belfast BT9 5PX
United Kingdom

Email: sally.watson@afbini.gov.uk

1   BACKGROUND

1.   The Technical Working Party on Automation and Computer Programs (TWC) have been asked to draft guidance on the use of Multiple Range Tests for Section 2 of TGP/8/1 Draft 7: PART II. This document is intended to focus attention on what the guidance should be. The topic of Multiple Range Tests and the closely-related topic of LSD's have been addressed below. As the use of such statistics to compare means in the examination of distinctness requires an appropriate underlying statistical model, a section on this topic has also been included.

2   LSD AND OTHER TESTS TO COMPARE MEANS

2.   The following sections give brief descriptions of the LSD and Multiple Range Tests, an explanation as to why the latter offer no advantages over the former in distinctness testing, and an explanation as to why an appropriate statistical model is needed in the comparison of means.

2.1   Least Significant Difference (LSD)

3.   The Least Significant Difference (LSD) is a statistic used to compare variety means from analysis of variance (ANOVA) of a characteristic and to make decisions about whether the varieties are significantly different from each other in that characteristic. In other words it represents the minimum difference between two variety means that the crop expert may declare to be different at a given significance level. The LSD is calculated using an estimate of random variation from the ANOVA.

4.   The LSD is not in itself a criterion used for making DUS decisions, but may form part of a criterion where, among other things, the nature of the variety means and the source of estimated random variation are specified. For example:-

- An LSD is used as part of the COYD distinctness criterion to compare over-year[1] variety means using the variety-by-year interaction as an estimate of random variation.

- An LSD is used as part of the 2x1% distinctness criterion to compare within-year variety means using the variety-by-replicate interaction as an estimate of random variation.

---

[1]   Although the term 'year' is interchangeable with the term 'independent growing cycle', in the following, for the sake of brevity, only the term year will be used.

▪ A form of LSD is used as part of the COYU uniformity criterion to compare the uniformity of the candidate variety with the uniformity of reference varieties.

5. It would be inconsistent with the rest of this document to describe the LSD in detail as descriptions can be found in many statistical text books, e.g. Clewer and Scarisbrick (2001). However, enough detail will be given to place it in context with the following sections on Multiple Range Tests and their comparison with LSD's.

6. The LSD is chosen to give a particular size of test (α%) when comparing two means using a single characteristic. This is also known as the significance level of the LSD, e.g. 5% or 1%. It means that if an LSD is used to make an *a priori* comparison, i.e. without knowledge of the data, then there is an α% chance of making a Type I error, i.e. declaring the two variety means to be significantly different when they are not different.

7. Although the LSD controls the comparison-wise Type I error chance, it does not control the experiment-wise Type I error chance. The more comparisons that an LSD is used to make, the greater the overall or experiment-wise Type I error chances. For example, if a 5% LSD is used to compare 14 independent pairs of means, then there is a 51% chance ($= 100\% \times (1 - (1 - 0.05)^{14})$) of declaring at least one of the pairs of variety means to be significantly different when they are not different.

8. The standard formula for an α% LSD to compare two means made up of *n* observations is:-

$$\text{LSD}\,\alpha\% = t_{(\alpha\%, rdf)} \times \sqrt{s^2 \times \frac{2}{n}}$$

Where s^2 is an estimate of random variation taken from the ANOVA, *rdf* is the degrees of freedom of s^2, and $t_{(\alpha\%,\ rdf)}$ is the two-sided α% critical value of the Student's t-statistic on *rdf* degrees of freedom.

2.2 <u>Multiple Range Test (MRT)</u>

9. A Multiple Range Test (MRT), also known as a multiple comparison test, is similar to an LSD in that it is:-

• a statistic used to compare variety means from analysis of variance (ANOVA) of a characteristic and decide about the significance of variety differences

• calculated using an estimate of random variation from the ANOVA

10. An MRT differs from an LSD in that it is:-

• chosen to give a particular size (α%) of test over all the comparisons for which it is intended for a characteristic, and so controls the experiment-wise Type I error chance. For this reason an α% MRT is larger than an α% LSD, and also the comparison-wise Type I error chance of an MRT is much smaller than its experiment-wise Type I error chance

• intended to be used to make *a posteriori* comparisons

11.   There are a number of different MRT's. The choice of which to use depends partly on the comparisons to be made: for example, if one particular variety mean is to be compared with all others, or if all variety means are to be compared with all others. Descriptions of MRT's can be found in many statistical text books, e.g. Clewer and Scarisbrick (2001).

2.3   The use of LSD rather than MRT in distinctness testing

12.   Distinctness testing involves a number of comparisons being made *a posteriori* - depending on which varieties are most similar. As an LSD is intended for *a priori* comparisons, the question may be raised as to why not use distinctness criteria based on an MRT, which is intended for *a posteriori* comparisons. The answer is that use of an MRT offers no advantages because in distinctness testing the focus is on comparing only certain pairs of varieties. The consequence of this is that the comparisons on which the experiment-wise Type I error chance of an MRT is based would not all be made, and the comparison-wise Type I error chances of the comparisons that are made would be inappropriately small compared to the comparison-wise Type I error chances from using an LSD.

2.4   Inferences and the underlying statistical model

13.   Exactly which source of variation modelled in the ANOVA is used as the estimate of random variation (s^2) in the comparison of means is very important in terms of the conclusions or inferences that can be drawn about the consistency of any differences between varieties that are declared to be significantly different. The source of variation might be

- the plant-to-plant variation within plots

- the replicate-to-replicate variation within varieties in a trial allowing for any block effects, i.e. the variety-by-replicate interaction

- the year-to-year variation within varieties allowing for overall year effects, i.e. the variety-by-year interaction

14.   In comparing the differences between variety means with an LSD or an MRT, the crop expert is effectively testing whether the difference in the variety means is larger than the difference that might reasonably have arisen due to chance or random variation affecting the observations making up the variety means when there was no difference between the varieties. Thus, it is critical that the estimate of random variation used in the LSD or MRT is appropriate to the conclusions that are sought. For example:-

- if the random variation is the variety-by-replicate interaction, then the crop expert can conclude that varieties with significantly different means are different relative to the variation from replicate to replicate within a variety.

- alternatively, if the random variation is the variety-by-year interaction, as in the COYD criterion, then the crop expert can conclude that varieties with significantly different means are different relative to the variation from year to year within varieties.

15.   This illustrates an important feature of the COYD criterion: implicit in its use is the test of whether variety differences are consistent over time. This is not the case in the use of an LSD or MRT within one year of trial, as the variation is from replicate to replicate in the trial and not over time. The 2x1% criterion somewhat avoids this problem: it incorporates a test of

consistency over time in its requirement that the varieties need to be significantly different in the same direction at the 1% level in at least two out of three years.

2.5   References

A. G. Clewer & D. H. Scarisbrick (2001). Practical Statistics and Experimental Design for Plant and Crop Science. John Wiley, Chichester.

[End of document]